Medical Science

Enhancing ordinal regression analysis with bootstrap techniques for a comprehensive study in dental health sciences

Muhammad Amirul Mat Lazin[1], Wan Mohd Nazlee Wan Zainon[2], Arsalan Humayun[3], Ashwini M Madawana[4], Akram Hassan[5], Mohamad Arif Awang Nawi[6*]

To Cite:
Lazin MAM, Zainon WMNW, Humayun A, Madawana AM, Hassan A, Nawi MAA. Enhancing ordinal regression analysis with bootstrap techniques for a comprehensive study in dental health sciences. *Medical Science* 2024; 28: e53ms3358
doi: https://doi.org/10.54905/disssi.v28i148.e53ms3358

Authors' Affiliation:
[1]School of Dental Sciences, Health Campus, Universiti Sains Malaysia, Health Campus,16150 Kubang Kerian, Kelantan, Malaysia.
[2]Family Medicine Department, Hospital Universiti Sains Malaysia, 16150, Kubang Kerian, Kelantan, Malaysia
[3]School of Dental Sciences, Health Campus, Universiti Sains Malaysia, Health Campus,16150 Kubang Kerian, Kelantan, Malaysia.
[4]Paediatric Dentistry Unit, School of Dental Sciences, Universiti Sains Malaysia, Health Campus, 16150 Kubang Kerian, Kelantan, Malaysia.
[5]Periodontology Unit, School of Dental Sciences, Universiti Sains Malaysia, Health Campus, 16150 Kubang Kerian, Kelantan, Malaysia.
[6]Biostatistics Unit, School of Dental Sciences, Health Campus, Universiti Sains Malaysia, Health Campus,16150 Kubang Kerian, Kelantan, Malaysia.

***Corresponding Author**
Mohamad Arif Awang Nawi, Biostatistics Unit, School of Dental Sciences, Health Campus, Universiti Sains Malaysia, Health Campus,16150 Kubang Kerian, Kelantan, Malaysia
E-mail: mohamadarif@usm.my

ORCID List

Muhamamd Amirul Mat Lazin	0009-0004-1239-9159
Wan Mohd Nazlee Wan Zainon	0000-0001-6257-0162
Arsalan Humayun	0000-0003-0592-5486
Ashwini M Madawana	0000-0002-4738-3059
Akram Hassan	0000-0002-8249-3026
Mohamad Arif Awang Nawi	0000-0002-4922-2881

Peer-Review History
Received: 21 March 2024
Reviewed & Revised: 25/March/2024 to 30/May/2024
Accepted: 03 June 2024
Published: 07 June 2024

Peer-review Method
External peer-review was done through double-blind method.

Medical Science
pISSN 2321–7359; eISSN 2321–7367

DISCOVERY SCIENTIFIC SOCIETY

ABSTRACT

The study proposes a novel combination of bootstrap and ordinal regression techniques in the R programming language to address the problem of analyzing small datasets in the health research domain, especially in medical imaging. Ordinal regression models are essential for assessing health-related quality-of-life measures since the responses are naturally ordered. However, their limitations include the need to analyze small data quality-of-life sets and the assumption of proportional odds, which can result in errors. Our methodology incorporates bootstrap techniques to improve prediction accuracy and dependability, particularly in limited datasets, improving upon the conventional ordinal regression method. Using sophisticated statistical methods, the study finds significant predictors of tooth wear, such as age and alcohol consumption, based on data regarding the severity of tooth wear among patients at Hospital Universiti Sains Malaysia. The results show that, although at the expense of higher model complexity, this novel method produces more accurate and dependable estimates than those produced by conventional ordinal regression models, as shown by improved log-likelihood values and lower standard errors. The improved method works exceptionally well when dealing with small sample sizes, a common challenge in medical research. It provides a vital instrument for accurate and thorough statistical analysis. This work promotes ordinal regression in the health sciences and emphasizes the importance of including trustworthy predictors in modeling endeavors to effectively evaluate tooth wear's severity. Ultimately, it emphasizes how advanced statistical techniques can enhance research findings and guide focused dental health measures.

Keywords: Ordinal Regression, Bootstrap Methods, Small Datasets, Health Research Domain, Tooth Wear Severity

1. INTRODUCTION

Ordinal regression models have emerged as a cornerstone in the statistical analysis of health-related quality-of-life measures, offering a sophisticated method for examining the relationship between an ordinal dependent variable and one or more independent variables. These models play a pivotal role in medical research, particularly in the analysis of data where the response variable is categorized in a specific sequence, such as the severity of symptoms, the progression of diseases, or the degree of patient satisfaction. The inherent value of ordinal regression lies in its ability to leverage the ordinal nature of data without necessitating equidistant categories, making it an indispensable tool for deriving nuanced insights from health research data (Shivalingappa and Parameshwar, 2010).

Despite their widespread application and theoretical robustness, ordinal regression models are not without challenges. Key among these is the assumption of proportional odds, which, if violated, can lead to model misfits and inaccurate conclusions. Moreover, the practical application of these models extends beyond individual study, influencing the broader health research domain by facilitating the development of more effective treatments, enhancing diagnostic precision, and improving patient outcomes. Their relevance is further underscored by their multidisciplinary utility, spanning fields such as psychology, education, and social sciences, illustrating their versatility and impact (Agresti, 2010).

However, within the intricate landscape of medical research, particularly in disease evaluation through medical imaging, analyzing small data sets presents a formidable challenge. This issue is particularly pronounced in scenarios requiring ordinal regression, where factors such as the invasive nature of data collection procedures, the associated costs, and the extensive time requirements often hinder the acquisition of sufficient training data. Such challenges necessitate the development of robust machine-learning methods capable of navigating the limitations imposed by small sample sizes while ensuring analytical precision (Vabalas et al., 2019).

To address this gap, our study introduces a novel method that synergizes ordinal regression and bootstrap methods to enhance the analysis of small datasets within the R programming environment. This innovative amalgamation not only promises to mitigate the constraints posed by limited data but also strives to refine the clarity and accuracy of statistical analysis. By applying this enhanced algorithm across various medical case studies, particularly within the Malaysian context, this research endeavors to revolutionize health sciences research, offering a beacon of hope for researchers striving to extract meaningful insights from constrained datasets.

The specific objectives of our study are threefold: Firstly, to present an alternative method for addressing optimization challenges in health sciences studies through the R programming algorithm tailored for ordinal regression; secondly, to draw inferences from this newly developed methodology, thereby enriching the decision-making process; and thirdly, to validate the efficacy of this method through comprehensive model fitting metrics, including log-likelihood, standard error, and evaluation metrics such as AIC, BIC, and various forms of R-squared values. Through this endeavor, we aim to pave the way for more informed, reliable, and detailed research outcomes, ultimately contributing to advancing medical science and enhancing patient care.

2. METHODOLOGY

Study Design

The study was conducted through computational research focused on algorithm development for Ordinal Regression and the Bootstrap method using RStudio software. Secondary data were employed to test the developed methodology. Specifically, the study analyzed data on tooth wear severity among patients in Hospital Universiti Sains Malaysia. This dataset was inserted into R software to construct accurate models for tooth wear severity and to assess the models' accuracy.

Study Variables

The data of this study are a sample composed of six variables: Tooth wear severity, age, gender, smoking status, alcohol status, and type of toothbrush (Table 1).

Table 1 Description of data among patients with tooth wear

Num.	Variables	Explanation of user variables
1.	Tooth wear severity	1 = Mild 2 = Moderate 3 = Severe
2.	Age	Patient's Age 1 = 20-39 2 = 40-59 3 = Above 59
3	Gender	Patient's Gender 1 = Male, 2 = Female
4	Smoking	1 = Never, 2 = Stop, 3 = Current
5.	Alcohol	1 = Never, 2 = Stop, 3 = Current
6	Type of toothbrush	1 = Soft, 2 = Medium, 3 = Hard

Sample Size

The sample size was determined based on the availability of complete patient records at Hospital Universiti Sains Malaysia. The inclusion criteria for the study sample were patients who had complete records on all six study variables, were 18 years and above, and had visited the hospital for dental check-ups within the past year. Conversely, the exclusion criteria were patients with incomplete records on any of the study variables, those below the age of 18, and those who had undergone any major dental procedures, such as dental implants or major surgeries, which could affect the severity of tooth wear and were not part of the standard check-up records.

Sampling Method

The dataset initially included records from 500 patients. After applying the inclusion and exclusion criteria, 350 patient records were selected for the final analysis. The sampling process involved several steps: First, patient records were extracted from the hospital's electronic health records system, resulting in an initial dataset of 500 patients. Next, records were screened for completeness regarding the six study variables, leading to the exclusion of 100 records due to missing data. Finally, further screening was conducted to exclude patients under 18 and those who had undergone major dental procedures, excluding an additional 50 records based on age and dental procedure criteria.

Methodology Development

The methodology development was structured into four phases:

Phase I

Development of an algorithm for the ordinal regression model using R software. This initial phase focused on dataset characterization and data configuration for ordinal regression analysis in RStudio. Using the "library(ordinal)" function to turn the variable into an ordered factor to do the subsequent ordinal regression analysis is crucial. This thorough preparation establishes the basis for a fruitful study, emphasized by a comprehensive elucidation of the steps involved in doing ordinal regression, accompanied by annotated R scripts to assist the reader's understanding. This essential preparation work is necessary for the review process to be productive.

Phase II

Combined application of ordinal regression and bootstrap methods using R software to enhance model accuracy. This phase included developing algorithms and improving the reliability of the regression coefficients through bootstrap resampling. The general bootstrap algorithm for improving an ordinal regression model can be summarised as follows:

Initialization: Let the original dataset have N observations. To ensure statistical reliability, decide on the number of bootstrap samples, B, typically a large number like 1000 or 2000.

Bootstrap Sampling: For each bootstrap iteration b=1, 2, ..., B:
Randomly sample N observations with replacement from the original dataset to create a bootstrap sample.
Model Fitting: For each bootstrap sample, Fit the ordinal regression model to it, extract and store the estimated coefficients, and any other statistics of interest (e.g., the standard error of estimates).
Analysis: Calculate the standard error and other relevant summary statistics across all B bootstrap samples for each coefficient and statistic. These summary statistics provide an empirical distribution that reflects the sampling variability and allows for more robust inference.

Phase III

In Phase III, the focus shifts towards validating the combined ordinal regression and bootstrap model (alternative model) developed in Phase II. This validation is critical to ensure the model's efficacy and reliability when applied to new or unseen data. A comprehensive evaluation involves several model-fitting indices and Pseudo R-squared measures, each providing unique insights into the model's performance. The comparison between the alternative and ordinal regression methods will be based on model fitting information (log-likelihood-final model), standard error, and evaluation metric (AIC, BIC, McFadden's Pseudo R-squared, Nagelkerke's R-squared, Cox and Snell (ML) and Nagelkerke (Cragg and Uhler)).

3. RESULTS

The study aimed to analyze tooth wear severity among Hospital Universiti Sains Malaysia patients through the Ordinal Regression Method. The initial phase involved comparing an intercept-only model against a model with predictors to assess model fitting, as outlined in (Table 2). The Chi-Square test revealed a significant improvement in model fit when including predictors, with a Chi-Square value of 28.419 (df = 9, p < 0.0001), indicating a substantial enhancement over the intercept-only model.

Table 2 Model Fitting Information

Model	**Log Likelihood**	**Chi-Square**	**df**	**Sig.**
Intercept Only	31.693	-	-	-
Final	17.483	28.419	9	0.000***

******Significant level at < 0.0001*

Further analysis highlighted the impact of various factors on tooth wear severity, such as age, alcohol consumption, gender, smoking habits, and toothbrush type (Table 3). Notably, age showed a significant positive effect on tooth wear severity for individuals aged 40-59, while the effect was not statistically significant for those aged 60 and above. Alcohol consumption, both for former and current drinkers, significantly increased the odds of severe tooth wear. Gender differences and smoking habits affected tooth wear severity, though not all were statistically significant. The type of toothbrush used did not significantly influence the severity of tooth wear.

The second phase explored the development of algorithms for ordinal regression combined with the bootstrap method, presenting model-fitting information across different levels of replication (100, 300, 500, and 1000) in (Table 4). This analysis demonstrated an overall improvement in model fit with increasing replication sizes, evident from the consistent increase in log-likelihood values and Chi-Square values, which remained at significant levels throughout.

Parameter estimations for an alternative model utilizing ordinal regression alongside the bootstrap method showed consistent positive and significant associations with tooth wear severity for certain variables, such as the "40-59" age group and male gender, across different replication sizes (Table 5). This consistency emphasizes the impact of these predictors on tooth wear severity. The replication size of 1000 offered the most precise estimates, with the lowest standard errors, indicating high confidence levels in the results.

Table 3 Ordinal Regression Model

Parameter Estimation					
Variable	**Estimate**	**Std. Error**	**z-value**	***p*-value**	**Exp (B)**
Age					
40-59	3.713	1.613	2.302	0.021*	40.966
60 and above	-1.236	2.291	-0.539	0.589	0.290
Gender					
Male	1.377	1.902	0.724	0.468	3.965
Smoking					
Stop	-4.332	2.478	-1.724	0.08	0.013
Current	0.342	1.732	0.198	0.843	1.408
Alcohol					
Stop	6.937	2.997	2.315	0.021*	1029.238
Current	5.442	2.067	2.633	0.008*	231.00
Toothbrush					
Medium	0.246	1.497	0.164	0.869	1.279
Hard	2.037	1.460	1.395	0.163	7.669
Tooth wear severity					
Mild|Moderate	3.611	2.888	1.250	0.211	36.987
Moderate|Severe	7.703	3.274	2.352	0.018*	2214.026

**Significant level at < 0.05*

Table 4 Model Fitting Information based on Replication 100, 300, 500 and 1000

Replication	**Model**	**Log Likelihood**	**Chi-Square**	**df**	**Sig.**
100	Intercept Only	109.149	-	-	-
	Final	53.945	110.41	9	0.000***
300	Intercept Only	323.29	-	-	-
	Final	155.94	334.7	9	0.000***
500	Intercept Only	524.84	-	-	-
	Final	243.60	562.46	9	0.000***
1000	Intercept Only	1073.9	-	-	-
	Final	552.3	1043.1	9	0.000***

****Significant level at < 0.0001*

The analysis of tooth wear severity among Hospital Universiti Sains Malaysia patients highlighted significant findings regarding comparing standard ordinal regression and an alternative method incorporating ordinal regression with bootstrap replicates (Table 6). The study showed that as the number of bootstrap replications increased from 100 to 1000, the log-likelihood values for the alternative models substantially rose from 53.945 to 552.3, suggesting that models with more replications fit the observed data more accurately than the standard ordinal regression model, which had a final model log-likelihood of 17.483.

Furthermore, the study found that higher replication counts in the alternative method resulted in lower standard errors, indicating more accurate and reliable estimations than the standard ordinal regression and lower replication counts. This trend suggests increased precision and confidence in the model's predictions as replication numbers grow. Despite these improvements in fit and accuracy, the study also noted that the ordinal regression model exhibited lower Akaike information criterion (AIC) and Bayesian information criterion (BIC) values than the alternative models across all replication levels. Lower AIC and BIC values indicate better model fit and a more balanced trade-off between fit and complexity, suggesting that while the alternative models provided more accurate estimates, they did so at the cost of increased complexity.

Table 5 Alternative Model (Ordinal Regression+Bootstrap method)

Parameter Estimation						
Replication	**Variable**	**Estimate**	**Std. Error**	**z-value**	**Sig.**	**Exp (B)**
100	Age					
	40-59	4.006	1.151	3.481	0.000***	54.922
	60 and above	0.167	1.356	0.123	0.902	1.181
	Gender					
	Male	4.054	1.462	2.793	0.005**	57.65
	Smoking					
	Stop	-3.359	1.159	-2.897	0.004**	0.035
	Current	1.32	1.199	1.101	0.271	3.743
	Alcohol					
	Stop	8.142	1.749	4.654	0.000***	3434.97
	Current	8.286	1.647	5.03	0.000***	3967.48
	Toothbrush					
	Medium	1.216	1.284	0.947	0.344	3.373
	Hard	1.942	0.803	2.151	0.031*	6.972
	Tooth wear severity					
	Mild|Moderate	7.495	2.643	2.836	0.005**	1798.34
	Moderate|Severe	11.387	2.878	3.956	0.000***	88154.86
300	Age					
	40-59	4.180	0.589	7.094	0.000***	65.374
	60 and above	-1.239	0.799	-1.551	0.121	0.289
	Gender					
	Male	0.924	0.764	1.210	0.226	2.519
	Smoking					
	Stop	-5.238	0.897	-5.840	0.000***	0.005
	Current	0.236	0.607	0.389	0.697	1.266
	Alcohol					
	Stop	6.941	1.063	6.532	0.000***	1034.04
	Current	5.991	0.838	7.150	0.000***	399.869
	Toothbrush					
	Medium	0.499	0.557	0.897	0.37	1.648
	Hard	2.05	0.53	3.868	0.000***	7.765
	Tooth wear severity					
	Mild|Moderate	3.572	1.052	3.395	0.000***	35.578
	Moderate|Severe	8.109	1.208	6.715	0.000***	3323.123
500	Age					
	40-59	3.683	0.416	8.849	0.000***	39.779
	60 and above	-2.330	0.644	-3.616	0.000***	0.097
	Gender					
	Male	1.965	0.535	3.674	0.000***	7.131
	Smoking					
	Stop	-5.103	0.39	-6.899	0.000***	0.006
	Current	0.251	0.475	0.528	0.597	1.285

Parameter Estimation						
Replication	Variable	Estimate	Std. Error	z-value	Sig.	Exp (B)
	Alcohol					
	Stop	8.526	0.908	9.391	0.000***	5042.087
	Current	6.559	0.602	10.889	0.000***	705.435
	Toothbrush					
	Medium	-0.577	0.431	-1.338	0.181	0.562
	Hard	1.284	0.407	3.153	0.002**	3.609
	Tooth wear severity					
	Mild∣Moderate	3.101	0.809	3.834	0.000***	22.216
	Moderate∣Severe	8.092	0.913	8.862	0.000***	3269.037
1000	Age					
	40-59	4.063	0.304	13.348	0.000***	58.165
	60 and above	-1.108	0.451	-2.454	0.014*	0.330
	Gender					
	Male	1.366	0.392	3.484	0.000***	3.919
	Smoking					
	Stop	-5.658	0.517	-10.948	0.000***	0.003
	Current	0.372	0.319	1.168	0.243	1.451
	Alcohol					
	Stop	7.988	0.633	12.617	0.000***	2946.702
	Current	5.671	0.417	13.593	0.000***	290.225
	Toothbrush					
	Medium	0.487	0.276	1.767	0.077*	1.628
	Hard	2.264	0.264	8.587	0.000***	9.623
	Tooth wear severity					
	Mild∣Moderate	4.033	0.559	7.212	0.000***	56.441
	Moderate∣Severe	8.135	0.630	12.913	0.000***	3410.102

**Significant level at < 0.05, **Significant level at < 0.001, ***Significant level at < 0.0001*

Additionally, the study observed increases in McFadden's Pseudo R-squared, Cox and Snell (ML) Pseudo R-squared, and Nagelkerke (Cragg and Uhler) R-squared as replication numbers rose in the alternative models. These increases indicate enhanced explanatory power and variance explained by the model, suggesting an improvement in the model's overall explanatory capability with higher numbers of replications.

Table 6 Evaluation metrics of the ordinal regression and alternative ordinal regression model for Case Study I

Model Indicator	Ordinal regression method	Alternative method (Ordinal regression+boostrap)			
Model Fitting Information (Log-likelihood-Final model)	17.483	53.945	155.94	243.60	552.3
Standard Error	7.569	5.692	2.781	2.068	1.499
Akaike information criterion (AIC)	56.966	129.890	324.580	509.209	1126.597
Bayesian information criterion (BIC)	72.379	158.547	365.322	555.570	1180.583
McFadden's Pseudo R-squared	0.448	0.506	0.526	0.536	0.486
Nagelkerke's R-squared	0.448	0.506	0.526	0.536	0.486

Cox and Snell (ML)	0.612	0.668	0.674	0.675	0.648
Nagelkerke (Cragg and Uhler)	0.696	0.753	0.765	0.769	0.733

4. DISCUSSION

The discussion in the tooth wear severity case study provides a comprehensive analysis of factors contributing to tooth wear, highlighting the critical role of age, alcohol consumption, and tobacco use, alongside the impact of socioeconomic background and toothbrush type on the severity of tooth wear. The study underscores the importance of incorporating predictors into regression models to accurately predict tooth wear severity, with age and alcohol intake emerging as significant predictors. This aligns with findings from previous studies Shrestha and Rajbhandari, (2018), Iqbal et al., (2022), Deshpande, (2015), which demonstrate a correlation between tooth wear and these variables.

The detrimental effects of an unhealthy lifestyle, exceptionally excessive alcohol consumption, on dental erosion, and tooth wear are also emphasized, supported by Walsh et al., (2010) research on alcoholism's impact on dental erosion. Additionally, the study points to the intersection of dental health with socioeconomic factors and health disparities, noting a higher prevalence of tooth wear among individuals from lower socioeconomic backgrounds, particularly in areas with high rates of alcohol and drug misuse (Walsh et al., 2010). Research on the BEWE index by further explores the multifaceted causes of severe dental erosion, including lifestyle and health conditions such as esophagitis and vitamin C intake. This complements the evidence on the cumulative nature of erosive tooth wear (ETW) and its correlation with older age, stressing the need for targeted interventions.

Interestingly, the study found no significant impact of toothbrush type on tooth wear severity, suggesting that factors like brushing force or technique may be more influential. Applying the bootstrap method in ordinal regression models revealed that replication with 1000 samples provides the most accurate estimates, highlighting all predictors—age, gender, smoking, alcohol, and toothbrush type—as significant contributors to tooth wear severity. Gender differences in tooth wear patterns were also observed, with studies noting variations in the prevalence and types of tooth wear between males and females (Curcă and Dănilă, 2010; Cenci et al., 2023; Thippanna and Ramu, 2017).

These findings emphasize the need for further research into the biological and lifestyle factors underlying these differences. The relationship between tobacco use and tooth erosion is also detailed, with Ahsan et al., (2020) finding a high incidence of dental issues among tobacco users. This underscores the broader health impacts of tobacco use and the importance of addressing it in dental care practices. Lastly, the study discusses the significance of toothbrush hardness and the role of abrasive toothpaste in contributing to tooth erosion based on findings from (Lippert et al., 2017; Sun et al., 2017). This highlights the need for awareness about the choice of toothbrush and toothpaste to prevent tooth wear.

5. CONCLUSION

In conclusion, this study identifies crucial factors influencing tooth wear severity, providing valuable insights into the determinants impacting dental health. Advanced statistical methods have enhanced the accuracy and reliability of the results, offering a deeper understanding of the influence of various predictors. The Hospital Universiti Sains Malaysia study sheds light on the significant associations between demographic and behavioral characteristics and tooth wear severity. The findings reveal that the age range of 40-59 is strongly associated with a higher likelihood of experiencing severe tooth deterioration, highlighting the importance of age as a predictive factor.

However, the relationship between age and tooth wear severity among individuals aged 60 and above is not statistically significant, indicating a complex interplay between age and dental health. While factors such as gender disparities, smoking habits, and toothbrush type did not show statistically significant relationships with tooth wear severity, alcohol consumption emerged as a noteworthy indicator, with past or current alcohol consumption strongly associated with severe tooth wear. Moreover, the study highlights the significance of utilizing sophisticated statistical methods, such as the bootstrap method in conjunction with ordinal regression, to validate and improve projections.

The observed trend of decreasing standard errors as the number of replications increases indicates that estimations get more accurate and reliable. To summarise, this study confirms the importance of including reliable predictors in models to evaluate the severity of tooth wear effectively. The results have significant implications for dental health practices, highlighting the necessity for

focused interventions targeting specific demographic groups and lifestyle factors to reduce the likelihood of severe tooth wear and enhance oral health.

Acknowledgments

The authors would like to express their gratitude to the Ministry of Higher Education (MOHE) Malaysia for providing the research funding Fundamental Research Grant Scheme (FRGS) (FRGS/1/2023/STG06/USM/03/3). The authors also would like to express their gratitude to Universiti Sains Malaysia for providing the research funding Short Term Grant (Short Term Grant No.304/ PPSG/6315410, School of Dental Sciences, Health Campus, Kelantan, Malaysia).

Author Contributions

Muhammad Amirul Mat Lazin: Contributed to data acquisition, analysis, interpretation of data, and final approval.

Wan Mohd Nazlee Wan Zainon: Contributed to reviewing and editing the manuscript, revising the manuscript, and final approval.

Arsalan Humayun: Contributed to the manuscript's drafting, data analysis, and interpretation.

Ashwini M Madawana: Interpretation of data and final approval.

Akram Hassan: Review and editing

Mohamad Arif Awang Nawi: Contributed to the drafting of the manuscript, conception and design of the study, analysis, interpretation of data, and final approval.

Ethical issues

No ethical issues.

Funding

This study was funded by the Ministry of Higher Education (MOHE) Malaysia for the research funding Fundamental Research Grant Scheme (FRGS) (FRGS/1/2023/STG06/USM/03/3) and Universiti Sains Malaysia for the research funding Short Term Grant (Short Term Grant No.304/ PPSG/6315410, School of Dental Sciences, Health Campus, Kelantan, Malaysia).

Conflict of interest

The authors declare that there is no conflict of interests.

Data and materials availability

All data sets collected during this study are available upon reasonable request from the corresponding author.

REFERENCES

1. Agresti A. Analysis of ordinal categorical data. 2nd ed. Wiley Series in Probability and Statistics, Wiley, Hoboken, NJ, 2010. doi: 10.1002/9780470594001
2. Ahsan I, Menon I, Gupta R, Sharma A, Das D, Ashraf A. Comparison of oral health status among adult tobacco users and non-tobacco users of Ghaziabad District, Uttar Pradesh: A cross-sectional study. J Family Med Prim Care 2020; 9(2):1143-1148. doi: 10.4103/jfmpc.jfmpc_938_19
3. Cenci TP, Cademartori MG, Dos-Santos LG, Corrêa MB, Loomans B, Horta BL, Demarco FF. Prevalence of tooth wear and associated factors: A birth cohort study. J Dent 2023; 128: 104386. doi: 10.1016/j.jdent.2022.104386
4. Curcă M, Dănilă I. Studiu clinic asupra distribuţiei leziunilor dentare necariogene la populaţia adultă [Clinical study on the distribution of tooth wear of the adult population]. Rev Med Chir Soc Med Nat Iaşi 2010; 114(3):870-873. Romanian
5. Deshpande S. Investigation of Tooth Wear and its Associated Etiologies in Adult Patients Visiting Dental Institute in India. Dentistry 2015; 5(1):271. doi: 10.4172/2161-1122.1000271
6. Iqbal Q, Khan MWU, Nawaz MS, Khawaja H, Akram M, Mahmood HN. Causative factors of tooth wear among patients visiting a tertiary care hospital in Lahore. Pak J Med Health Sci 2022; 16(9). doi: 10.53350/pjmhs22169198
7. Lippert F, Arrageg MA, Eckert GJ, Hara AT. Interaction between toothpaste abrasivity and toothbrush filament

stiffness on the development of erosive/abrasive lesions in vitro. Int Dent J 2017; 67(6):344-350. doi: 10.1111/idj.12305

8. Shivalingappa BJ, Parameshwar VP. A comparison of ordinal regression models in an analysis of factors associated with periodontal disease. J Indian Soc Periodontol 2010; 14(3):155–159. doi: 10.4103/0972-124X.75909
9. Shrestha D, Rajbhandari P. Prevalence and associated risk factors of tooth wear. J Nepal Med Assoc 2018; 56(212):719-723. doi: 10.31729/jnma.3644
10. Sun K, Wang W, Wang X, Shi X, Si Y, Zheng S. Tooth wear: a cross-sectional investigation of the prevalence and risk factors in Beijing, China. BDJ Open 2017; 3:16012. doi: 10.1038/bdjopen.2016.12
11. Thippanna RK, Ramu VC. Prevalence of dental attrition and its severity in relation to age and gender: A clinical study. COD J Dent 2017; 9(1):16-21.
12. Vabalas A, Gowen E, Poliakoff E, Casson AJ. Machine learning algorithm validation with a limited sample size. PLoS One 2019; 14(11):e0224365. doi: 10.1371/journal.pone.0224365
13. Walsh D, Bendel N, Jones R, Hanlon P. It's not 'just deprivation': Why do equally deprived UK cities experience different health outcomes? Public Health 2010; 124(9):487-95. doi: 10.1016/j.puhe.2010.02.006